WORKING PAPER
DEPARTMENT
OF ECONOMICS

SEQUENTIAL INNOVATION, PATENTS, AND IMITATION

James Bessen
Eric Maskin

No. 00-01 January 2000

MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

50 MEMORIAL DRIVE
CAMBRIDGE, MASS. 02142

1\16

Sequential Innovation, Patents, and Imitation*

by

James Bessen and Eric Maskin

Working Paper
11/99

Abstract: How could such industries as software, semiconductors, and computers have been so innovative despite historically weak patent protection? We argue that if innovation is both sequential and complementary—as it certainly has been in those industries—competition can increase firms' future profits thus offsetting short-term dissipation of rents. A simple model also shows that in such a dynamic industry, patent protection may reduce overall innovation and social welfare. The natural experiment that occurred when patent protection was extended to software in the 1980's provides a test of this model. Standard arguments would predict that R&D intensity and productivity should have increased among patenting firms. Consistent with our model, however, these increases did not occur. Other evidence supporting our model includes a distinctive pattern of cross-licensing in these industries and a positive relationship between rates of innovation and firm entry.

James Bessen
Research on Innovation
jbessen@researchoninnovation.org

Eric Maskin
Harvard University and Massachusetts Institute of Technology
emaskin@mit.edu

*This research was supported in part by a grant from the NSF.

1. Introduction

The standard economic rationale for patents is to protect potential innovators from imitation and thereby give them the incentive to incur the cost of innovation. Conventional wisdom holds that, unless would-be competitors are constrained from imitating an invention, the inventor may not reap enough profit to cover that cost. Thus, even if the social benefit of invention exceeds the cost, the potential innovator without patent protection may decide against innovating altogether.[1]

Yet, interestingly, some of the most innovative industries today—software, computers, and semiconductors—have historically had weak patent protection and have experienced rapid imitation of their products. [2] Defenders of patents may counter that, had stronger intellectual property protection been available, these industries would have been even more dynamic. But we will argue that the evidence suggests otherwise.

In fact, the software industry in the United States was subjected to a revealing natural experiment in the 1980's. Through a sequence of court decisions, patent protection for computer programs was significantly strengthened. We will show that, far from unleashing a flurry of new innovative activity, these stronger property rights ushered in a period of stagnant, if not declining, R&D among those industries and firms that patented most.

We maintain, furthermore, that there was nothing paradoxical about this outcome. For industries like software or computers, there is actually good reason to believe that imitation *promotes* innovation and that strong patents (long patents of broad scope) *inhibit* it. Society might be well

[1] This is not the only justification for patents, as we will discuss in Section 2. But it is the rationale most commonly advanced. Also, patents provide a spillover benefit through their disclosure requirements. We focus here, as does the traditional literature, on the capacity of patents to block competitors.

[2]Software was routinely considered excluded from patent protection until court decisions in the late 80's. Semiconductor and computer patent enforcement was very uneven until the organization of the Federal Circuit court in 1982. Both areas contend with substantial problems of prior art [Aharonian, OTA, 1992], and some experts contend that as many as 90% of semiconductor patents are not truly novel and therefore invalid [Taylor and Silberston, 1973]. These problems make consistent enforcement difficult.

Surveys of managers in semiconductors and computers consistently report that patents only weakly protect and incent innovation. Levin et al [1987] found that patents were rated weak at protecting the returns to innovation, far behind the protection gained through lead time and learning curve advantages. Patents in electronics industries were estimated to increase imitation costs by only 7% [Mansfield, Schwartz and Wagner] and 7 - 15% [Levin, et al]. Taylor and Silberston [1973] found that very little R&D was performed to take advantage of patent rights.

As one might expect, diffusion and imitation are rampant in these industries. Firms formally and informally exchange information themselves, employees frequently move from one firm to another, and spin-off firms are common. More important, imitation lags are brief. Tilton [1971] estimated that the time from initial discovery to commercial imitation in Japan averaged just over one year in semiconductors in the 60's. In software, imitation lags are sometimes shorter.

served if such industries had only limited intellectual property protection. Moreover, many firms might genuinely *welcome* competition and the prospect of being imitated.[3]

This is because these are industries in which innovation is both *sequential* and *complementary.* By "sequential," we mean that each successive invention builds on the preceding one—in the way that Windows built on DOS. And by "complementary," we mean that each potential innovator takes a somewhat different research line and thereby enhances the overall probability that a particular goal is reached within a given time.[4] Undoubtedly, the many different approaches taken to voice-recognition software, for example, hastened the availability of commercially viable packages.

The sequential and complementary nature of innovation is widely recognized, especially in high tech industries. Gort and Klepper [1982] found an average of 19 sequential improvements to 23 major innovations and many more uncounted improvements. Analysis of many innovations has found that most of the productivity gain is achieved via improvements to the original innovation (see for example, Enos [1962]). Scotchmer [1996], Green and Scotchmer [1995], and Chang [1995] have studied the theoretical implications of sequential innovation. A variety of empirical studies have found strong evidence of innovative complementarities, including Jaffe [1986] and Henderson and Cockburn [1996] (see a review of the literature in Griliches [1992]). Spence [1984] and Levin and Reiss [1988] have studied the effects of spillover complementarities in a non-sequential environment.

A firm that patents its product in a world of sequential and complementary innovation can prevent its competitors from using that product (or sufficiently similar ideas) to develop further innovations. And because these competitors may have valuable ideas not available to the original firm about how to achieve such innovations, the patent may therefore slow down the pace of invention.

Of course, patent traditionalists have a counter-argument to this train of logic too: if a jealously guarded patent seriously interferes with innovative activity, the patent-holder can simply license it (thereby allowing innovation to occur). After all, presumably he could capture the value added

[3] When IBM announced its first personal computer in 1981, Apple Computer, then the leading personal computer-maker, responded with full-page newspaper ads headlined "Welcome IBM. Seriously." A high tech cliché is that competition "expands the market."

of such innovation by an appropriately chosen licensing fee (or so the argument goes). The flaw in this argument is that licensing creates competition and, as we will show in Section 3, the profit-dissipating effect of this competition may well outweigh the additional profit created by greater innovation. In such a case, the patent-holder would choose not to license, even though society would suffer from that decision.

But whether or not patent protection is available, a firm may well be better off if other firms imitate and compete with it. Although imitation reduces the firm's current profit, it raises the probability of further innovation and thereby improves the prospect that this firm will make another profitable discovery later on.

In short, when innovation is sequential and complementary, standard reasoning about patents and imitation may get turned on its head. Imitation becomes a *spur* to innovation, while strong patents become an *impediment.*

We will proceed as follows. In Section 2 we will review the static model that underlies the traditional justification for patents. We show that, besides helping to ensure that innovative firms cover their costs, patents can also encourage innovative activity on the part of firms that otherwise would be more inclined to imitation. Then in Section 3 we turn to a dynamic model and delineate the circumstances in which patents inhibit innovation and firms are made better off by being imitated. Finally, in Section 4 we discuss the evidence for the pertinence of this dynamic model to high tech industries.

2. The static model

We consider an industry consisting of two (*ex ante* symmetric) firms.[5] Each firm can undertake R&D to discover and develop an innovation with expected (social) value v. The cost of R&D is c and, if a single firm incurs that cost, the probability of successful innovation is p.[6]

---

[4] One example of complementarity treated in the literature is that of information spillovers, following Arrow [1962]. But we do not limit ourselves to this example.

[5] Limiting the model to two firms is a matter of expositional convenience only; all our results extend to three or more firms. Indeed, we will see below that our arguments about the inefficiency of patents and the drawback of mergers as a corrective to that inefficiency become stronger with more than two firms.

We assume that if a firm is not copied, it can capture the social value of its innovation.[7] However, the other firm, unless prevented by patent protection, can costlessly develop an imitation. And, if it does so, the innovating firm obtains only a fraction s ($s < \frac{1}{2}$) of the value of v (to simplify matters, we assume that the imitating firm captures this same fraction s).[8]

We suppose that the expected social value of undertaking R&D is positive:

(1) $$p \cdot v - c > 0.$$

However, the net payoff from R&D to a firm that anticipates being imitated is only $s \cdot p \cdot v - c$. And, even if (1) holds, we may well have

(2) $$s \cdot p \cdot v - c < 0.$$

The combination of (1) and (2) represents the classic incentive failure that the patent system is meant to address. Formula (2) captures the idea that, without patent protection, a firm cannot make money on R&D investment when its potential innovation is imitable. A patent proscribes imitation and, therefore, guarantees an innovator the full net social return on R&D expenditure ($p \cdot v - c$). Formula (1) then tells us that, as long as society gains from R&D investment, so will the innovator himself.

But even in a setting where (2) does *not* hold—so that R&D remains profitable despite imitation—patents may well serve a useful purpose. This is because they can encourage several firms to go after the same innovation. Typically, different firms have different ideas about how to solve a particular technological problem. Therefore, increasing the number of firms in pursuit of a solution raises the probability that *someone* will succeed; this is what we called "complementarity" in the introduction.

---

[6] Our setting in this section is static but, if it is viewed as the reduced form of a dynamic model, p could also be interpreted as the discount factor corresponding to the time lag to innovation.

[7] This is, of course, a strong assumption. However, the incentive failure and inefficiency from monopoly that arise when it is not satisfied are already well understood. The assumption is a simple way to abstract from these familiar distortions.

[8] The assumptions that imitation is costless and that an imitator enjoys the same share of profit as the innovator (i.e., the innovator has no first-mover advantage) are, of course, unrealistic. By invoking them, however, we are strengthening the case for patents. This will bolster our argument in Section 3 where we point out the shortcomings of the patent system.

We model the complementarity by assuming that, if both firms undertake R&D, the event that firm 1 is successful is *statistically independent* of that of firm 2.[9] Hence, the total probability of successful innovation with two firms investing is

$$1-(1-p)^2 = 2p - p^2,$$

and the social net benefit of this investment is

$$(2p - p^2)v - 2c. \tag{3}$$

Formula (3) implies that the marginal social benefit of the second firm's R&D expenditure is $(2p - p^2)v - 2c - (pv-c)$. Let us suppose this is positive, i.e.,

$$(p - p^2)v - c > 0. \tag{4}$$

Now, in the absence of patent protection, a firm's expected profit if both firms undertake R&D is

$$s(2p - p^2)v - c \tag{5}$$

($2p- p^2$ is the probability of successful innovation, after which each firm enjoys a payoff of sv). But even if (5) is positive (and (4) implies that it will be, provided that s is not too much smaller than ½), an equilibrium in which both firms invest may not be sustainable. Indeed, if one firm undertakes R&D, the other will do so too only if (5) exceeds

$$p \cdot s \cdot v, \tag{6}$$

the expected payoff from free-riding on the innovative firm's R&D activity. That is, even when (5) is positive, only one firm will invest if

$$s(p - p^2)v - c < 0. \tag{7}$$

With the prospect of patent protection, by contrast, a firm undertaking R&D can expect a payoff of

---

[9] More realistically, the techniques available to each firm might be correlated to some degree, leading to correlation between the two

(8) $$\tfrac{1}{2}(2p - p^2)v - c,$$

if the other invests as well (we assume that even if both firms make the discovery, only one obtains the patent,[10] and that each has an equal chance of this). Note that (8) is positive as long as (4) holds. Thus, if it is socially desirable for a second firm to invest, patent protection will induce it to do so, whereas, without such protection, it might well merely imitate. Patents accomplish more, therefore, than merely protecting innovators from imitation. Indeed, they create a risk of *over*-investment: (8) could be positive even if (4) fails to hold.[11]

We can summarize our observations in the following result (which is also illustrated in Figure 1):

> Proposition 1: In the above (static) model, patent protection gives rise to at least as much R&D investment[12] (and hence innovation) as in either (a) a regime without patent protection or (b) the social optimum.

Observe that the possible over-investment in R&D induced by patents could, in principle, be avoided if there were no complementarities of research across firms. Specifically, one could imagine awarding a firm an *ex ante* patent, e.g., the right to develop a vaccine against a particular disease. [13] Such protection would, of course, serve to deter additional firms from attempting to develop the innovation in question. And this would be efficient, provided that these other firms would not enhance the probability or speed of development, i.e., provided that they conferred no complementarity.

Notice that there is no room for patent licensing in this static model. A patent-holder derives a payoff v. Were he to license to his competitor, the total industry payoff would be $2sv$, which is less than v.

---

firms' chances of success. We assume statistical independence (no correlation) only for convenience.

[10] This gets at the idea that patents have *breadth*, and so a patent-holder can hold up the implementation of other firms' discoveries that are similar, but not identical, to his own.

[11] The possibility that patents can give rise to excessive spending on R&D is well known from the patent-race literature; see Dasgupta and Stiglitz [1980] and Loury [1979].

[12] And possibly strictly more.

[13] Wright [1983] similarly explores patent awards and government contracts.

This simple static model thus captures basic results of patent race models such as Loury [1979] and Dasgupta and Stiglitz [1980]. It also captures aspects of static models that consider spillover complementarities such as Spence [1984], which emphasizes the socially redundant R&D that can occur under patents. These results, however, change sharply when dynamic considerations are introduced.

3. Dynamic model

Let us now enrich the model to accommodate a *sequence* of potential innovations, each of which builds on its immediate predecessor. More specifically, we will suppose that, for a firm to have a realistic chance of developing the innovation of the next generation, it must have market experience with that of the current generation. That is, it must be operating at the "state of the art." The idea behind this assumption is that any new product or process would spring directly from those of the current generation, and so "hands on" experience with the latter is a prerequisite for innovation. For example, the firms surveyed by Levin et al [1987] (in more than one hundred manufacturing industries) reported that typically only a few other firms are capable of duplicating their innovations and that learning curves, lead time and sales and service efforts provide significant obstacles to imitation.

Formally, consider an infinite sequence of potential innovations, each of which makes the previous generation obsolete and has incremental value v over that generation. Given that the generation t innovation has already been developed, a firm that incurs R&D cost c and participates in the generation t market has positive probability of discovering the generation $t+1$ innovation. If there is just one firm attempting to innovate, this probability is p. If this firm undertakes R&D in every generation, the expected number of innovations is

$$(9) \qquad p(1-p) + 2p^2(1-p) + 3p^3(1-p) + \ldots = \frac{p}{1-p},$$

where the first term on the left-hand side of (9) is the expectation of one innovation, the second term is the expectation of two, etc. If an innovation fails to occur in any given generation, no further innovation is possible. The firm's overall expected profit and the social net benefit of R&D are both given by

(10) $$-c + p(1-p)(v-c) + 2p^2(1-p)(v-c)\ldots = -c + \frac{p}{1-p}(v-c) = \frac{p \cdot v - c}{1-p},$$

since we continue to assume that a monopolist can capture all the social surplus.

If, for each innovation, there are two firms investing in R&D and participating in the market, then, given the same complementarity as in the static model, the expected number of innovations is

$$(2p - p^2)(1-p)^2 + 2(2p - p^2)^2(1-p)^2 + \ldots = \frac{2p - p^2}{(1-p)^2},$$

and the social surplus contributed by the second firm is

(11) $$\frac{(2p - p^2)v - 2c}{(1-p)^2} - \frac{pv - c}{1-p} = \frac{pv - (1+p)c}{(1-p)^2}.$$

Let us retain the assumption that, if one firm imitates the other's innovation, each enjoys a share s of the total gross value. We first explore the nature of equilibrium when there is no patent protection.

If one firm attempts to innovate in each generation, while the other only imitates, the innovator's overall expected profit is

(12) $$\frac{psv - c}{1-p}.$$

However, the other firm will choose to imitate only if

(13) $$\frac{s(2p - p^2)v - c}{(1-p)^2} < \frac{psv}{1-p}.$$

The left-hand side of (13) is the firm's expected profit when both firms undertake R&D in every generation. The right-hand side is its profit if it always imitates. Formula (13) simplifies to

(14) $$p \cdot s \cdot v < c.$$[14]

Observe that (12) and (14) imply that *neither* firm will invest if (14) holds, and that *both* will invest if it fails to hold.

Comparing (14) with its static-model counterpart (7), we see that there is a greater incentive for a second firm to invest in the dynamic setting than in the static model (notice that this is true only for the second firm; the incentives for R&D are exactly the same in either model for the first firm). This is because the second firm's R&D in the dynamic model raises the probability, not only of the next innovation, but of subsequent innovations, and this is advantageous to the firm even if, subsequently, it merely imitates the first firm.

Next we consider what happens under a patent system, assuming that each successive innovation is patentable and, for simplicity, each patent applies only to a single innovation. Although subsequent innovations are not covered by a given patent, each patent still allows patent holders to block entry to the subsequent markets—firms must participate in the original market in order to succeed with a follow-on product. That is, a patent conveys a hold-up right over subsequent innovations.

With patents, at least one firm is willing to invest provided that (10) is positive, i.e.,

(15) $$\frac{v}{c} > \frac{1}{p}.$$

Indeed, both firms will be willing to "race" to the first innovation (with perhaps the loser then dropping out altogether) provided that

$$-c + \tfrac{1}{2}(2p-p^2)\frac{v-c}{1-p} > 0, \text{ i.e.,}$$

(16) $$\frac{v}{c} > \frac{2-p^2}{2p-p^2}.$$

---

[14] We would obtain the same formula if we compared the left-hand side of (13) with the profit from one-time imitation followed by R&D investment in subsequent generations.

But from (11), it is not efficient for both firms to invest in R&D if

$$(17) \qquad \frac{1+p}{p} > \frac{v}{c}.$$

And since the left-hand side of (17) exceeds the right-hand side of (16), we obtain, for this first-generation innovation, the standard result of excessive R&D.

Matters are quite different, however, in subsequent generations. Indeed, both firms will continue to undertake R&D after the first generation only if patent-holders are prepared to license their innovations. The patent-holder can find licensing profitable only if the joint profits of both firms exceed the monopoly profits the patent-holder could obtain without licensing. But this might not be so. The total value created will be larger with licensing because of complementarities, but competition may dissipate a share of this value to consumers, resulting in lower joint profits.

Let us assume that licensing fees take the form of lump-sum payments.[15] In this case, product market competition would be unaffected by the patent and license. Thus, ignoring fees, each firm would earn sv, and so a licensee would expect to earn

$$(18) \qquad sv + \frac{s(2p-p^2)v-c}{(1-p)^2},$$

where the first term in (18) is the current profit from competing against the patent-holder and the second term is expected future profit (assuming that all subsequent innovations are licensed, so that *both* firms can profitably undertake R&D). Total joint profits are

$$(19) \qquad 2sv + \frac{2s(2p-p^2)v-2c}{(1-p)^2} = \frac{2sv-2c}{(1-p)^2}$$

and the patent holder could extract this amount by charging a fee equal to (18).

Thus, the license-holder will license if and only if

[15] This assumption would make sense, say, if the licensor could not readily monitor the extent to which the licensee's output made use of the licensed item. For example, it may not be clear how important a patented idea is for the source code of a software product. This would seriously interfere with setting fees as a function of output. Even if the licensor *could* monitor the licensee's output, it might well not be able to monitor the licensee's *costs*, in which case it would optimally set a fee that with positive probability was higher than the licensee would accept (in the event that the licensee's costs turned out to be high). In this setting with uncertainty about costs, we would obtain the same conclusion as below: too little licensing and hence too little innovation.

(20) $$\frac{v-c}{1-p} < 2sv + \frac{2s(2p-p^2)v-2c}{(1-p)^2},$$

where the left-hand side of (20) is (monopoly) profit without licensing and the right-hand side is duopoly profit plus the licensing fee. Formula (20) can be rewritten as

(21) $$\frac{v}{c} > \frac{1+p}{2s-1+p}.$$

Thus unlike the static model, licensing could be to a patent-holder's advantage in this dynamic setting.

But for s sufficiently small, we have

(22) $$\frac{1+p}{p} < \frac{1}{sp} < \frac{1+p}{2s-1+p},$$ [16]

where, from (11), $\frac{1+p}{p}$ is the smallest value of $\frac{v}{c}$ for which it is efficient for both firms to invest, and, from (14), $\frac{1}{sp}$ is the smallest value for which both firms will *choose* to invest in the absence of a patent system. Then innovations in the range $\frac{1}{sp} < \frac{v}{c} < \frac{1+p}{2s-1+p}$ would be pursued in a market without patents, but would not be licensed with patents. Thus, whether or not patents are available, the market will engender too little R&D. But, with sufficiently vigorous competition (s sufficiently small), (22) implies that a market with patents will, in general, be less efficient than one without, because patent-holders will not license sufficiently often (and so the pace of innovation will be even slower than in an economy without patents).

We have stressed the case where (21) fails, and so patent licensing does not suffice to ensure adequate innovation. But even when (21) holds—the case in which innovations are sufficiently "important"—there is still an important difference between the static and dynamic models.

---

[16] The left-hand inequality in (22) always holds; the right-hand inequality holds if $s < \frac{1}{2+p}$.

Whereas a firm would invariably shun competition and imitation by the other firms in the former setting, it may welcome them in the latter. To see this, notice that in the static model, a firm expects payoff $pv-c$ without competition, but at most $\frac{1}{2}(2p-p^2)v-c$ (the payoff would be even lower were patents not available) when it has a rival. Hence, competition always makes a firm worse off. In the dynamic mode, however, a lone firm's payoff is given by (10), whereas, when (21) holds, its expected payoff, whether or not patents are available, is

$$\frac{s(2p-p^2)v-c}{(1-p)^2},$$

which is greater than (10). Hence, the firm's profit is actually enhanced by imitation and competition in this case, because of the greater rate of innovation that they foster. We summarize our findings (also illustrated in Figure 2) as follows:

> Proposition 2: After the first-generation innovation, patent protection gives rise to efficient R&D (and the absence of patent protection gives rise to insufficient R&D) if and only if it is socially optimal for just one firm to invest. When having more than one firm undertake R&D is efficient, however, a regime without patents induces an R&D investment level (and hence a pace of innovation) that (although still too low) is, in general, more efficient than one with patent protection (provided that competition is sufficiently intense). Moreover, if innovations are sufficiently important (v/c is sufficiently big), not only the social but the private return to R&D (i.e., a firm's profit) is enhanced by competition and imitation.

Remark: Although it is true that patent protection is efficient in the case where it is socially desirable to have only one firm invest, the "specialness" of this case is perhaps masked by our assumption that there are only two firms. Had we assumed half a dozen firms instead, the case in which exactly one innovating firm is optimal would appear much more restrictive. Moreover, this case also is presumably most relevant for innovations that are only "marginal." For "important" innovations (high returns), society would want multiple firms to be in the race. Finally, our assumptions of costless imitation and no first-mover advantage (see footnote 8) also stack the deck in favor of patents. With more realistic assumptions, patents might not confer an advantage even when one firm is efficient.

In our model, regimes with and without patents both give rise to too little innovation. One may ask, therefore, whether there is some superior third alternative. In fact, in our model a *merger* between the two competing firms would presumably permit their complementary ideas to flourish and of course would eliminate cutthroat rivalry altogether. Thus, the merger provides a theoretical solution to the R&D incentive problem. However, it is not likely to do so in practice, although it is often attempted in the software and computer industries. Indeed, the conclusion in our model that merger is optimal rests heavily on our simplifying (and heroic) assumption that monopoly creates no distortion. Here again our two-firm model may be misleading. In a more realistic model with a larger number of rivals, merger of all of them would become correspondingly more difficult and distortionary.

4. Empirical evidence of dynamic innovation

We present three pieces of evidence suggesting that the dynamic rather than the static model applies to innovative industries: 1.) The pattern of cross-licensing in high tech industries, 2.) The positive correlation between rates of innovation and rates of firm entry, and, 3.) The natural experiment in patent protection of software.

Cross-licensing of entire patent portfolios

The distinctive nature of patent licensing in several high tech industries is difficult to reconcile with traditional intellectual property models. Under static models of innovation, firms offer patent licenses to competitors only under restrictive conditions. Specifically, in the static model presented here, patent holders would never offer licenses to competing firms. However, this model assumes a "drastic" product innovation. Other models that allow for "non-drastic" process improvements find that under some product market conditions, sufficiently minor innovations would be offered to competing firms. None of these models, however, countenance situations where firms license their *entire* patent portfolios for a given field to *direct* competitors.

Yet cross-licenses covering entire fields were common between competitors during the first several decades of the semiconductor and computer industries. In the semiconductor industry during 1975 alone, 34 cross-licensing (and second source) agreements between rival firms were announced [Webbink, p. 99]. These licenses covered whole portfolios of patents related to an

entire technical field rather than individual inventions, and many covered future as well as existing patents.

Moreover, these cross-licenses were not intended as barriers to entry as has sometimes been the case in other industries [Scherer, 1980, see also Bittlingmayer, 1988]. A survey of company executives concluded "...all company officials interviewed agreed that no company had ever been prevented from entering the semiconductor business because of patents, nor had any company ever been refused a patent license [Webbink, p. 100]."

Indeed, this cross-licensing is fully consistent with the dynamic model of innovation. [17] In the first instance, the cross-licenses relieve competing firms of holdup problems. The dynamic model suggests that patents provide initial inventors with holdup rights on subsequent inventions. Survey evidence from the semiconductor industry [Hall and Ham, 1999] and from all industries [Cohen, et al., 1998] indicates that firms view these holdup rights as a major feature of patents. Under the static model, a patent could also holdup other innovations, if, for instance, patent grants were excessively broad. Although overly broad patents and uncertain enforcement may exacerbate holdup problems, industry participants are clear that typical semiconductor innovations draw on hundreds of previous developments and that it is primarily sequential innovation that generates opportunities for holdup [Grindley and Teece, 1997, Hall and Ham, 1999].

For this reason, high tech cross-licensing displays a dynamic concern with possible future benefits that is at odds with the static model. For example, although an established firm's portfolio might have been sufficient to prevent the entry of a startup competitor, most established firms apparently recognized that startups might contribute significantly to future industry innovations. Hence according to Webbink's survey of semiconductor firms, cross-licenses were readily offered to new firms (often at very modest royalties), contrary to the static model.

---

[17] Fershtman and Kamien [1992] model cross-licensing for situations where two complementary technologies are required to make a product. The logic is similar to the one suggested by our dynamic model in that firms make potential (but unknown and therefore uncontractable) complementary contributions to future products. The difference is that Fershtman and Kamien describe the exchange of specific *patents*, whereas in the semiconductor industry, consistent with the dynamic model, firms exchanged whole *portfolios* covering an entire field.

Moreover, licensing contracts typically cover future developments.[18] For example, when Robert Noyce founded Intel in the late 60's, he obtained cross-licenses from Texas Instruments, IBM and Fairchild that gave them access to future patents to the year 1999. The typical semiconductor cross-license today provides access to patents developed five years into the future. In many cases, the licenses provide use of covered patents for the entire term of the patent (now 20 years) [Grindley and Teece, 1997].

This dynamic concern for future developments guided licensing behavior from the beginning of the semiconductor industry. If ever an innovative firm could have gathered large monopoly rents from its patents, Bell Laboratories could have done so from its basic transistor patents. Instead, Western Electric (in charge of licensing the patents) aggressively offered the patents to all comers at a low 2% royalty beginning in 1953, dropping this royalty entirely in 1956.[19] The vice president for electronic component development explained the logic:

> We realized that if this thing [the transistor] was as big as we thought, we couldn't keep it to ourselves and we couldn't make all the technical contributions. It was to our interest to spread it around. If you cast your bread on the water, sometimes it comes back angel food cake [Tilton, p75-6].

Such behavior illustrates a strong concern with potential future benefits that goes well beyond static considerations of holdup. When different firms make complementary and sequential technical contributions (the dynamic model), collaborative forms of licensing may be privately advantageous.

During the 1980's, industry norms began to erode substantially as some established firms chose to milk their patent portfolios for royalties rather than use them for new developments. Also, changes in the legal regime made it easier for patent holders to sue and win, most notably the organization of the Federal Circuit court for patent appeals in 1983 [Angel, 1994]. This diminished cross-licensing, especially to new firms, e.g., Intel has sued sub-licensees of its original cross-licenses. But these developments do not alter the significance of semiconductor cross-licensing as an example of firms acting on dynamic concerns about future innovation.

---

[18] Note that in our dynamic model a market with an agreement that all future patents will be licensed is equivalent to one in which there are no patents at all.

Innovation and firm entry

The relationship between innovation and firm entry provides another test of the two models.

In the static model of intellectual property, innovation incentives depend on the patent holder's ability to extract monopoly rents. The magnitude of these rents depends on product market conditions. Rents will be greatest when the patent holder enjoys a complete monopoly; rents will generally be inferior when other firms can enter the product market, even if they produce only imperfect substitutes.[20] Thus a high rate of firm entry is often taken as *prima facie* evidence of insufficient appropriability.

One important source of variation in firm entry occurs over the product life cycle of an innovation and is related to patent protection. From studies by Gort and Klepper [1982], we know that entry conditions change over the life of major innovations. Initially, a single firm will typically enjoy a monopoly, often supported by an initial set of patents. When these patents expire (or entry is allowed for other reasons), firms freely enter. Over time additional innovations are made to improve the product. As some firms build large portfolios of in-force patents on these improvements, entry once again becomes more difficult. Combined with maturing demand, new firms stop entering and less successful firms exit, resulting eventually in a stable industry.[21]

If the static model holds, innovation rates should be greatest when entry is most limited and innovation should decline when large numbers of new firms enter. Also, markets that experience greater entry should, in general, exhibit lower rates of innovation. By contrast, if the dynamic model holds, innovation rates might well be greater during phases of high entry and also for products experiencing high entry. Entrants may bring complementary knowledge that increases industry prospects for successful innovation.

The evidence squarely supports the dynamic model. Gort and Klepper assembled information on innovation rates and entry rates for 23 major new products from the transistor to the zipper. They

---

[19] After 1956 Western Electric agreed to offer these patents royalty-free as part of an antitrust consent decree. However, industry observers and AT&T executives have commented that AT&T was willing to do this in any case.

[20] For example, consider the case where an innovator holds a patent on an improvement to a base product. The rents on the improvement will be greatest when the innovating firm has a monopoly on the base product as well. In general, the innovating firm will not realize the same rents from the improvement if entrants can freely offer an unimproved base product as a substitute—the unimproved version of the product can be offered at a lower price, tending to dissipate some of the rents.

[21] This exposition highlights the role of patents in changing conditions for entry; Gort and Klepper provide a broader explanation.

divide these data into five phases of each product's life-cycle, the phases defined by net entry behavior. The first phase is the monopoly stage (or near-monopoly in a few cases), the second exhibits positive net entry, in the third, entrants roughly balance exiters, the fourth has negative net entry, and the fifth exhibits rough stability again. For each phase they report the annual rate of entry and of innovation. Innovation rates are further divided into rates for major innovations and for minor innovations.

Table 1 and Figure 3 show the means (weighted by duration) of the annual rates of innovation for each phase for both major innovations and total innovations. As can be seen, neither the initial monopoly phase nor the final phases—both periods when entry is most constrained—have particularly high rates of innovation. The highest rates of innovation appear instead during the second and third phases, during and immediately following the period of greatest firm entry.

This result can be explored more formally with a Poisson model of the innovation count data. For the *i*th product during a phase of duration Δt_i, we assume that the hazard for an innovation, λ_i, is an exponential function of the net entry rate of firms, n_i:

$$\lambda_i = \Delta t_i \cdot e^{\alpha + \beta n_i}.$$

The probability that the number of innovations during this period is y is

$$P(Y=y) = \frac{e^{-\lambda_i} \cdot \lambda_i^y}{y!}.$$

It is possible that changes in both the rate of innovations and the rate of firm entry result from exogenous changes in technological opportunities. That is, firms may choose to enter when opportunities to innovate are greater. In this case, the independent variable would be correlated with the error term. To correct for this, we perform an instrumental variables estimation.

We begin, however, with a straightforward maximum likelihood estimation of this simple model displayed in column 1 of Table 2, both for all innovations (top) and for only major innovations (bottom). The results show a significant positive relationship between entry and innovation.

The Poisson regression model assumes that the variance of y equals the mean. But this will not be the case if there are stochastic errors in addition to the Poisson sampling error [see Hausman,

Hall and Griliches, 1984, Cameron and Trivedi, 1986]. To allow for possible "over-dispersion," we also performed the negative binomial regression described in Cameron and Trivedi. These regressions are shown in column 2. The coefficients are quite similar, positive and highly significant. A likelihood ratio test indicates that over-dispersion does exist, hence the negative binomial model is preferred.

For comparison we also perform a nonlinear least squares estimation in column 3. This method is consistent, but not efficient for our model, and so we estimate standard errors using White's heteroscedastic-consistent method [1980]. Results are similar, but in the estimation on major innovations the coefficient for entry is significant only at the 5% level.

To correct for possible endogeneity in the independent variable we instrument the rate of firm entry with two variables. Desirable instruments should be correlated with the rate of entry, but uncorrelated with changes in technological opportunity. The first instrument is the average number of firms in the industry over the entire product life.[22] Although technological opportunity may influence the *gross* rate of entry, the exit process is independent, and the equilibrium number of firms over all phases would seem to be determined by market size and structure independently of opportunity. The second instrument is a simple dummy flag that takes the value of 1 during the initial monopoly phase and 0 otherwise. The initial monopoly period, if it exists, is presumed to result from an original set of strong, broad patents and should thus be independent of technological opportunity as well.[23]

Estimates using these instruments are shown in Column 4. Again, coefficients are similar and the coefficient on firm entry is significantly positive.[24]

---

[22] These supplementary data were graciously provided by Steven Klepper.

[23] If initial monopolies do not arise from patent protection, then the static model would be irrelevant in any case. Note further that since we are instrumenting a nonlinear least squares estimation, the instruments apply to the pseudo-regressors of a linearized model, not to the rate of entry directly. To correspond to the form of the pseudo-regressors, the instruments were multiplied by the duration of the phase. Also, terms were included using the square of the average number of firms, the phase duration and a constant.

[24] A Hausman specification test could not reject the null hypothesis that the simple nonlinear least squares estimator was consistent. The instrumental variables estimation was repeated using only the second instrument, the initial patent flag, plus phase duration and a constant. Results were positive with an even higher coefficient. Generally similar although sometimes less significant results were also obtained including industry dummies and performing a fixed effects analysis conditioning on the product sums [Hausman, Hall and Griliches, 1984].

The marginal effect of firm entry is not large—roughly 30 or 40 entrant firms correspond to one innovation. But this does not alter the interpretation. Firm entry does not decrease innovation as suggested by the static model; instead, new firms increase the rate of innovation.

The Natural Economic Experiment in Software

The semiconductor, computer and software industries have historically experienced high levels of innovation despite weak patent protection. This suggests that the dynamic model is applicable, but, by itself, this evidence is not conclusive. Although these industries have been innovative without strong patent protection, perhaps they would have been *far more innovative* with strong protection; perhaps these industries offer many technological possibilities, but only the most highly profitable possibilities are realized under weak patent protection.

Fortunately, this alternative explanation can be tested. The patent courts subjected the software industry to a natural economic experiment during the 1980's.[25] Before this time, patent protection for innovations was very limited; instead, innovations were protected by copyright. This meant practically that direct copying of a software product was prohibited, but that copying the ideas and concepts embodied in software was not. Market entry therefore required significant investment in development, but entry could not be barred.

A series of court decisions in the early 1980's had the effect of extending patent protection to many software ideas. Consequently the number of patents issued annually covering software grew exponentially from the mid-80's to about 7,000 in 1995 (see Figure 4). Within the software industry, this has sometimes been described as a case of "fixing what ain't broke." Advocates counter, arguing along the lines of the static model, that increased patent protection should increase software innovativeness [USPTO, 1994].

If the static model is correct, then the extension of patent protection should have produced a sharp increase in R&D spending among those firms and industries applying for patents. This should have subsequently been followed by an increase in productivity growth. The changes should be measurable and large after controlling for other, possibly offsetting changes.

---

[25] Some other natural experiments involving the extension of patent protection are [Scherer and Weisburst, 1995] and [Challu, 1995].

According to the static model, R&D should increase with patent protection because firms can profitably pursue R&D projects that yield smaller returns, that is, projects with lower values of *v/c*. This can be seen as follows. As shown in Figure 1, projects with low values of *v/c*, labeled "insufficient innovation" under no patents, become feasible for a greater number of firms with the extension of patent protection. Assume a stationary distribution of R&D opportunities ranked by *v/c* such that $F(v/c)$ is the cumulative R&D spending required to invest in all opportunities with a return less than *v/c*. For simplicity assume *F* is concave. For the case of no patents, designate the entry threshold value of *v/c* for one firm as T_1^n and the threshold for two firms as T_2^n (the values of these are given in Figure 1). Then all opportunities with returns between *x* and $x+dx$ such that $T_1^n \le x < T_2^n$ will consume in total $dF(x)$ R&D dollars and will generate expected value of $p \cdot x \cdot dF(x)$. Opportunities where $T_2^n \le x$ will require $2dF(x)$ R&D dollars (two firms investing) and will generate expected value of $p(2-p) \cdot x \cdot dF(x)$. Then the average value of *v/c* for the industry is the ratio of total value *v* to total R&D investment,

$$A_n = \frac{p \cdot \int_{T_1^n}^{T_2^n} x \cdot dF(x) \ + \ p(2-p) \cdot \int_{T_2^n}^{\infty} x \cdot dF(x)}{1 \cdot \int_{T_1^n}^{T_2^n} dF(x) \ + \ 2 \cdot \int_{T_2^n}^{\infty} dF(x)} = \frac{p \int_{T_1^n}^{\infty} x \cdot dF(x) \ + \ p(1-p) \int_{T_2^n}^{\infty} x \cdot dF(x)}{\int_{T_1^n}^{\infty} dF(x) \ + \ \int_{T_2^n}^{\infty} dF(x)}.$$

With patents, the equivalent thresholds are T_1^p and T_2^p and the corresponding average value of *v/c* is A_p. Now examining the table in Figure 1, it is true that $T_1^p < T_2^p < T_1^n < T_2^n$. Using this, it is straightforward to show that $A_p < A_n$.

In other words, the *average* value of *v/c* should decrease for industries with the extension of patent protection. This logic can be readily extended to cases with more than two firms. Also, allowing each firm to have an equal chance of being an early-mover for any R&D opportunity means that the average value of *v/c* should also decrease for firms, or alternatively, the average value of *c/v* should increase.

For empirical analysis it is useful to note two aspects of this predicted change. First, since productivity is increasing in these industries (see below), the net social value *v* will increase at

least as fast as output. Therefore, an increase in c/v implies an increase in the ratio of R&D spending to output. In other words, the extension of patents should cause an increase in *relative* (to output) R&D spending. Relative R&D spending is a more useful measure than absolute R&D spending given the changing composition of industries as firms acquire, divest, startup and discontinue product lines and industries grow.

Second, v represents a discounted stream of future values. Typically the increase in value (and the associated increase in output) associated with an innovation will follow the expenditure of R&D only after considerable delay. For this reason, we should expect the ratio of R&D to output to increase quite rapidly upon the extension of patent protection and subsequently level off.

In contrast, if these industries follow the dynamic model, then we should expect the imposition of patents to lead to a reduction of R&D and productivity growth. This reduction should arise as patent holders stake out exclusive claims in different research areas inhibiting complementary innovation.

Unlike the static model, however, we should expect *this* transition to be quite gradual for two reasons. First, any patent holder faces a large body of well-established prior art and possibly competing claims. In such an environment, a patent portfolio capable of fencing off an area of research can be built up only gradually. In fact, there has been relatively little software patent litigation so far and the companies with large portfolios are only just beginning to pursue software patent claims [Business Week, 1997].[26]

Second, some of the most innovative firms may be reluctant to aggressively pursue patent claims. As we have seen above, although static firms will be better off with patent protection, dynamic firms may actually be better off without it. Thus the most innovative firms might seek to maintain industry norms of cooperation rather than to aggressively exert all patent rights. In this case, these norms will deteriorate slowly, and so problems of exclusive development may appear slowly. In fact, support for cooperative norms appears strong among many innovative software companies—senior executives from companies such as Microsoft, Sun and Oracle have expressed a general reluctance to pursue patent litigation and view their patenting activity as primarily defensive [PC Magazine, 1997, USPTO, 1994].

Indeed, pure software companies as a whole have not applied for many patents. A naive view might expect software patents to be obtained predominately by firms in the computer programming and data processing industry (SIC 737). In fact, the largest software patentees are in the computer hardware and telecommunications industries—industries which sell software products and also incorporate software in hardware products. The top 10 U. S. firms obtaining software patents in 1995 are listed in Table 3. The top ranked pure software firm in 1995 was Microsoft (rank 24) with 39 software patents.[27]

To summarize, if the static model holds, relative R&D spending should have increased sharply, followed by productivity. If the dynamic model applies, relative R&D spending and productivity would not have increased and might exhibit a mild decrease.

We examine these changes among three different samples of firms:

1.) The top 10 U.S. software patentees in 1995, accounting for 35% of the software patents issued to U.S. companies in that year,

2.) The industry groupings for computer hardware and programming services in the NSF R&D survey (company R&D funds for SIC 357 and part 737 and 871) [NSF, 1996, 1997], and,

3.) The grouping of computer, telecommunications and electronic components (SIC 357, 365-7) in the NBER R&D Masterfile [Hall, 1988], a listing of publicly traded U.S. firms.[28]

For the first and last samples, the R&D and sales measures are global. For the NSF sample, the R&D measures are domestic only and we measure R&D intensity using the NSF figures for sales for SIC 357 and 737.[29]

---

[26] Also, given the incompleteness of patent portfolios, much of this activity is directed not toward exclusive control of a market, but toward extracting royalties. Nevertheless, excessive royalties may limit complementary activity at the margin.

[27] The software patent series used in this analysis were developed by Greg Aharonian of the Internet Patent News Service. The criteria for software patents include not only the USTPO patent class, but also detailed examination of the specification, claims and abstract.

[28] Data for the top 10 firms was obtained from annual reports, 10-Ks and the NBER Masterfile. The series for AT&T was based on consolidated figures including NCR, the computer company which was purchased by AT&T in 1991. For this reason we use only the top 9 firms prior to 1991, although the difference is not significant. Both the NSF samples and the NBER R&D Masterfile are firm-based surveys where all of the R&D and sales of a firm are assigned to the SIC category of the firm's major product line. Thus our measures are diluted by non-software R&D and non-software output. Nevertheless, as long as software development constitutes a substantial portion of R&D, then we should expect to see a significant increase in R&D intensity.

[29] Note that beginning in 1985 FASB required that a portion of software development expense should be capitalized, hence reported R&D includes directly expensed items plus the amortization expense of capitalized software. The introduction of this change may

We initially explore R&D spending relative to sales (R&D intensity) rather than output. The trend in these measures is shown in Figure 5. The late 80's display a leveling off and possibly a reversal of an upward trend in research intensity over the previous decade. There does not appear to be so much as a 10% increase in R&D intensity among the firms and industries obtaining software patents.[30]

There could be two sorts of offsetting changes: 1.) Technological opportunities may have simultaneously fallen abruptly, and, 2.) The cost of performing R&D could have simultaneously risen sharply. A decline in technological opportunity seems at odds with the continued rapid growth and rapid innovation in these industries. Bronwyn Hall [1993] performs an econometric analysis on the same NBER dataset and finds that the output elasticity of R&D did not fall during the 1980's, but instead rose.[31]

Hall also presents evidence that the general costs of performing R&D did not rise sharply. If R&D costs had increased overall, offsetting an erstwhile increase in R&D spending, then the R&D intensity of other industries should have *fallen*. Figure 6 presents ratios of R&D intensity of software-related industries to the R&D intensity of the entire manufacturing sector. As can be seen, the relative R&D intensity of software-related industries fell over this time period. Thus not only did these industries fail to show a large increase in relative R&D spending, but they lagged behind the rest of the manufacturing sector over this period.

It is possible, however, that R&D spending relative to sales may understate R&D relative to output because of price effects. That is, as firms gain monopoly power with patents, prices may rise, inflating the sales figure in the denominator. To consider this possibility, Figure 7 displays the ratio of real R&D to output where R&D has been deflated using the NBER R&D deflator and sales have been deflated by a shipments-weighted index derived from the NBER Productivity Database for the industries involved. As can be seen, R&D relative to output exhibits a

---

have had a slight distortionary effect on reported R&D, tending to delay a portion of expenditures. The effect of this accounting change was to spread the impact of any sharp changes in R&D spending over two or three years. This effect was temporary, significant largely for pure software firms and of relatively brief duration (software is typically amortized over three years or less). This was not a substantial factor for the 10 largest software patentees and, based on this, would not seem to be a major factor for industry measures either.

[30] The NSF series becomes erratic after 1992 as the result of sample changes and as some firms were re-classified into different industries.

[31] The increase was concentrated among smaller public firms as large firms apparently lost productivity switching from mainframe technology to microcomputers. But overall technological opportunity did not decline.

significant decline during the late 1980's. Perhaps prices have been mis-measured for the computer industry. However, it seems unlikely that any measurement error could be so large as to mask major price increases. Hence this evidence is hard to reconcile with the static model.

Bronwyn Hall has suggested [1993] that competition may have hit the large mainframe firms in the industry especially hard as new firms entered the computer industry in the early 1980's. Consequently the response of the large firms (and by implication industry averages) might not be representative of firms in the industry as a whole. To consider this possibility, we examined two sub-samples from the software-related firms in the NBER R&D Masterfile: a balanced panel of 49 small firms and an unbalanced panel of new public firms.[32] Figure 8 shows the R&D intensity of these panels compared to the performance of the top 9 software patentees. As can be seen, R&D spending diverged between these groups during the early 1980's, consistent with Hall's interpretation, but these groups did not increase relative R&D spending either during the late 1980's in response to software patents.

Thus the extension of patent protection to software did not generate a relative increase in R&D spending as predicted by the static model; instead, consistent with the dynamic model, R&D spending seems to have remained roughly steady or to have declined. Not surprisingly, these industries did not demonstrate increased productivity growth as a result of the patent bonanza, as seen in Figure 9. Although multi-factor productivity may have fallen for reasons related to the transition from mainframes to microcomputers, there is no evidence of any underlying productivity increase commensurate with the increase in patents.

Of course, these industries have remained innovative and productivity growth is still positive. This does not, however, contradict the dynamic model; rather, the negative effects of the patent extension may not be felt for some time as industry cooperative norms continue and as litigation remains limited. The bill for this experiment has not yet come due.

In summary, the initial high level of innovation in software, mixed industry support for patents, and an apparent gradual slowdown of R&D intensity all suggest the applicability of the dynamic model. This conclusion is also supported by the distinctive pattern of cross-licensing in the

---

[32] The small firms are all those existing in 1980 and 1990 with fewer than 1,000 employees in 1980. The new firms are defined as firms that first enter the NBER R&D Masterfile after 1973 and have fewer than 5,000 employees their first recorded year. Conversations with Compustat confirmed that this procedure was likely to screen out most spin-offs, re-organizations and listing changes. New firms were dropped from the panel after 8 years.

semiconductor and computer industries and by the more general positive relationship between innovation and firm entry. All of this evidence is difficult to reconcile with the traditional static model of intellectual property.

Conclusion

Intellectual property appears to be one of those areas where results that seem secure in the context of a static model are overturned in a dynamic model. Imitation invariably inhibits innovation in a static world; in a dynamic world, imitators can provide benefit to both the original innovator and to society as a whole. Patents preserve innovation incentives in a static world; in a dynamic world, firms may have plenty of incentive to innovate without patents and patents may constrict complementary innovation.

This suggests a cautionary note regarding intellectual property protection. The reflexive view that "stronger is always better" is incorrect; rather a balanced approach is required. The ideal patent policy limits "knock-off" imitation, but allows developers who make similar, but potentially valuable complementary contributions. In this sense, copyright protection for software programs (which has gone through its own evolution over the last decade) may have achieved a better balance than patent protection. In particular, industry participants complain that software patents have been too broad and too obvious, leading to holdup problems [USTPO]. Also in this regard, patent systems that limit patent breadth, such as the Japanese system, may offer a better balance. Thus our model suggests another, different rationale for narrow patent breadth than the recent economic literature on this subject.

Bibliography

AHARONIAN, G. 1995. *Patent News Service.*

ANGEL, DAVID P. 1994. *Restructuring for Innovation: the remaking of the US semiconductor industry,* , Guilford, NY.

ARROW, K. 1962. "Economic Welfare and the Allocation of Resources for Inventions," *The Rate and Direction of Inventive Activity,* Princeton Univ.

BARTELSMAN, E. AND GRAY, W., 1996. "The NBER Manufacturing Productivity Database", *NBER Technical Working Paper,* no. 205.

BITTLINGMAYER, G. 1988. "Property rights, progress, and the aircraft patent agreement," *Journal of Law and Economics,* v. **31**, p. 227.

CAMERON, A. AND TRIVEDI, P. 1986. "Econometric models based on count data: comparisons and applications of some estimators and tests," *Journal of Applied Econometrics,* v. **1**, p. 29.

CHALLU, P. 1995. "Effects of the monopolistic patenting of medicine in Italy since 1978," *International Journal of Technology Management*, **v. 10**, p. 237.

CHANG, H. 1995. "Patent scope, antitrust policy and cumulative innovation," *RAND Journal of Economics*, **v. 26, no.1**, p. 34.

COHEN, W., GOTO, A., NAGATA, A, NELSON, R., AND WALSH, J. 1998. "R&D Spillovers, Patents and the Incentives to Innovate in Japan and the United States," unpublished working paper.

DASGUPTA, P. AND MASKIN, E. 1987. "The simple economics of research portfolios," *The Economic Journal*, **v. 97**, p. 581.

DASGUPTA, P. AND STIGLITZ, J. 1980. "Uncertainty, Industrial Structure, and the Speed of R&D," *Bell Journal of Economics*, **11**, p. 1.

ENOS, J. 1962. *Petroleum Progress and Profits: a history of process innovation*, MIT Press..

FERSHTMAN, C. AND KAMIEN, M. 1992. "Cross licensing of complementary technologies," *International Journal of Industrial Organization*, **v. 10**, p. 329.

GREEN, J. AND SCOTCHMER, S. 1995. "On the division of profit in sequential innovation," *RAND Journal of Economics*, **v. 26, no.1**, p. 20.

GORT, M. AND KLEPPER, S. 1982. "Time paths in the diffusion of product innovations," *The Economic Journal*, **v. 92**, p. 630.

GRILICHES, Z. 1992. "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, **v. 94, supplement**, p. 29.

GRINDLEY, P. AND TEECE, D. 1997. "Managing intellectual capital: licensing and cross-licensing in semiconductors and electronics," *California Management Review*, **v. 39, no. 2**, p. 8.

HALL, B. 1988. "The manufacturing sector master file: 1959-1987," *NBER Working Paper Series*, **No. 3366**, p. .

HALL, B. 1993. "Industrial research during the 1980s: Did the rate of return fall?," *Brookings Papers: Microeconomics 2*, , p. 289.

HALL, B. AND HAM, R. 1999. "The patent paradox revisited: determinants of patenting in the US semiconductor industry, 1980-94," *NBER Working Paper Seris*, **no. 7062**.

HAUSMAN, J., HALL, B., AND GRILICHES, Z. 1984. "Econometric models for count data with an application to the patents-R&D relationship," *Econometrica*, **v. 52, no. 4**, p. 909.

HENDERSON, R. AND COCKBURN, I. 1996. "Scale, scope and spillovers: the determinants of research productivity in drug discovery," *RAND Journal of Economics*, **v. 27, no. 1**, p. 32.

JAFFE, A. 1986. "Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits and market value," *American Economic Review*, **v. 76, no. 5**, p. 984.

LEE, T. AND WILDE, L. 1980. "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economic*, **94**, p. 429.

LEVIN, R. 1982. "The Semiconductor Industry," *Government and Technical Progress: A cross-industry analysis*, **Nelson, R. ed.**, Pergamon, NY.

LEVIN, R. AND REISS, P. 1988. "Cost-reducing and demand-creating R&D with Spillovers," *RAND Journal of Economics*, **v. 19, no. 4**, p. 538.

LEVIN, R., KLEVORICK, A., NELSON, R. AND WINTER, S. 1987. "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity [3]*, , p. 783.

LOURY, G. 1979. "Market Structure and Innovation," *Quarterly Journal of Economic*, **93**, p. 395.

MANSFIELD, E., SCHWARTZ, M. AND WAGNER, S. 1981. "Imitation Costs and Patents: An empirical study," *Economic Journal*, **v. 91**, p. 907.

MYHRVOLD, N. 1997. "Interview," *PC Magazine*, **Sept. 9, 97**, p. 30.

NATIONAL SCIENCE FOUNDATION. Annual. *Research and Development in Industry.*

NATIONAL SCIENCE FOUNDATION. Annual. *Science and Engineering Indicators.*

NORDHAUS, W. 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*, MIT, Cambridge.

OFFICE OF TECHNOLOGY ASSESSMENT 1992. *Finding a Balance: computer software, intellectual property and the challenge of technological change*, Congress of the US.

SAGER, I. 1997. "Big Blue is out to collar software scofflaws," *Business Week*, **March 17, 1997.**

SCHERER, F. 1980. *Industrial market structure and economic performance*, , Rand McNally.

SCHERER, F. AND WEISBURST, S. 1995. "Economic effects of strengthening pharmaceutical patent protection in Italy," *International Review of Industrial Property and Copyright Law*, p. 1009.

SCOTCHMER, S. 1996. "Protecting early innovators: should second-generation products be patentable?," *RAND Journal of Economics*, **v. 27, no. 2,** p. 322.

SPENCE, M. 1984. "Cost reduction, competition, and industry performance," *Econometrica*, **v. 52, no. 1,** p. 101.

TAYLOR, C. T. AND SILBERSTON, Z. A. 1973. *The economic impact of the patent system: a study of the British experience*, Cambridge Univ. p. 294.

TILTON, J. 1971. *International Diffusion of Technology: The Case of Semiconductors*, , Brookings Inst.

USPTO 1994. "Hearings on Software Patent Protection," *United States Patent and Trademark Office*, **January-February 1994.**

WEBBINK, D. 1977. "The semiconductor industry : a survey of structure, conduct, and performance : staff report to the Federal Trade Commission," *Federal Trade Commission*.

WHITE, H. 1980. "A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedacticity," *Econometrica*, **v. 48,** p. 817.

Table 1 . Weighted means of annual rates of innovation by phase.

	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Rate of all innovations	0.39	0.57	0.62	0.36	0.43
Rate of major innovations	0.19	0.29	0.22	0.18	0.22
Rate of net firm entry	0.22	5.05	-0.07	-4.97	0.16

Source: Gort and Klepper, 1982

Table 2. Regressions on innovation counts.

Regression	1	2	3	4
Regression method	Poisson	Negative Binomial	Nonlinear least squares	Nonlinear least squares instrumental variables
Dependent variable	Total number of innovations			
Coefficient of firm net entry rate	.046* (.007)	.034* (.003)	.052* (.014)	.062* (.009)
Constant	-.820* (.052)	-.694* (.105)	-.934* (.366)	-.943* (.120)
θ in NEGBIN II	--	1.373* (.216)	--	--
R^2	.43	.37	.44	.43
Dependent variable	Number of major innovations			
Coefficient of firm net entry rate	.041* (.011)	.035* (.003)	.045 (.022)	.046* (.015)
Constant	-1.547* (.074)	-1.444* (.105)	-1.668* (.261)	-1.666* (.132)
θ in NEGBIN II	--	1.560* (.216)	--	--
R^2	.35	.30	.36	.36

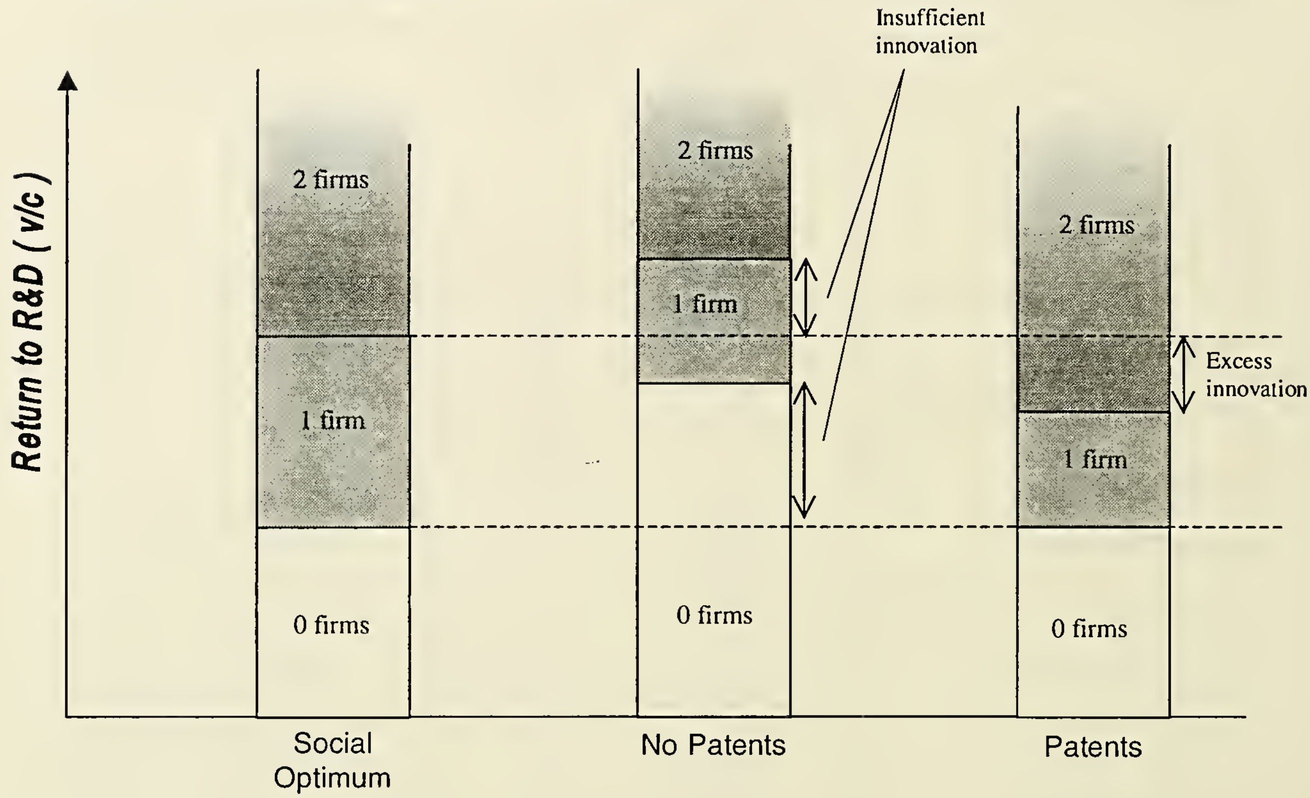
*Significant at the 1% level.
Asymptotic standard errors in parentheses, using White heteroscedasticity consistent standard errors for nonlinear regressions. Regressions cover 418 total innovations, 200 of these rated as major innovations, during 77 product phases over 887 product-years. Data are from Gort and Klepper [1982]. Instruments include a flag indicating initial monopoly phase, the average number of firms over the entire product life cycle, the square of the average number of firms, all multiplied by the duration of the phase, phase duration, and a constant. A Hausman specification test between the third and fourth columns does not reject the null hypothesis that the third column is consistent (P = .346 for all innovations and P = .917 for major innovations). NLLS-IV using only the monopoly flag times phase duration, duration and a constant as instruments generates significant and even larger coefficients on net entry.

Table 3. Top 10 Software Patentees, 1995

Firm	Software Patents Issued 1995	Total Utility Patents Issued 1995	R&D Spending 1994 (millions)
International Business Machines	503	1383	$3,382
AT&T	185	638	$3,110
Motorola	157	1012	$1,860
Xerox (including Fuji Xerox)	121	551	$895
Hewlett Packard	89	470	$2,027
Digital Equipment	80	189	$1,301
General Electric	59	758	$1,176
Apple Computer	57	129	$564
Ford Motor Co.	53	334	$5,214
Eastman Kodak	49	772	$859

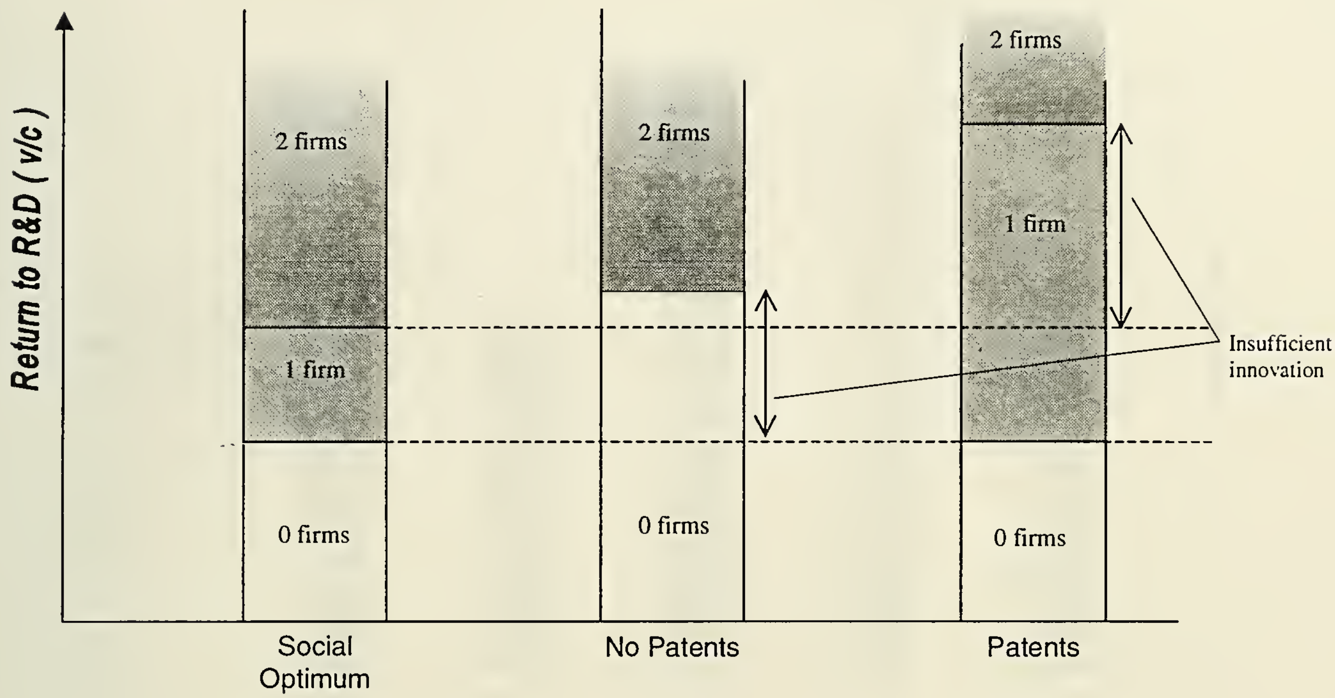
Sources: PATNEWS, USPTO, annual reports

Figure 1.Number of firms innovating in static model.



Innovating firms	Social Optimum	No Patents	Patents
One firm	$\frac{v}{c} > \frac{1}{p}$	$\frac{v}{c} > \frac{1}{ps}$	$\frac{v}{c} > \frac{1}{p}$
Two firms	$\frac{v}{c} > \frac{1}{p(1-p)}$	$\frac{v}{c} > \frac{1}{ps(1-p)}$	$\frac{v}{c} > \frac{1}{p(1-p/2)}$

Figure 2. Number of innovating firms (after first stage) in dynamic model.



Innovating firms	Social Optimum	No Patents	Patents
One firm	$\frac{v}{c} > \frac{1}{p}$	--	$\frac{v}{c} > \frac{1}{p}$
Two firms	$\frac{v}{c} > \frac{1+p}{p}$	$\frac{v}{c} > \frac{1}{ps}$	$\frac{v}{c} > \frac{1+p}{2s-1+p}$

Figure 3. Innovation rates during product life-cycle phases

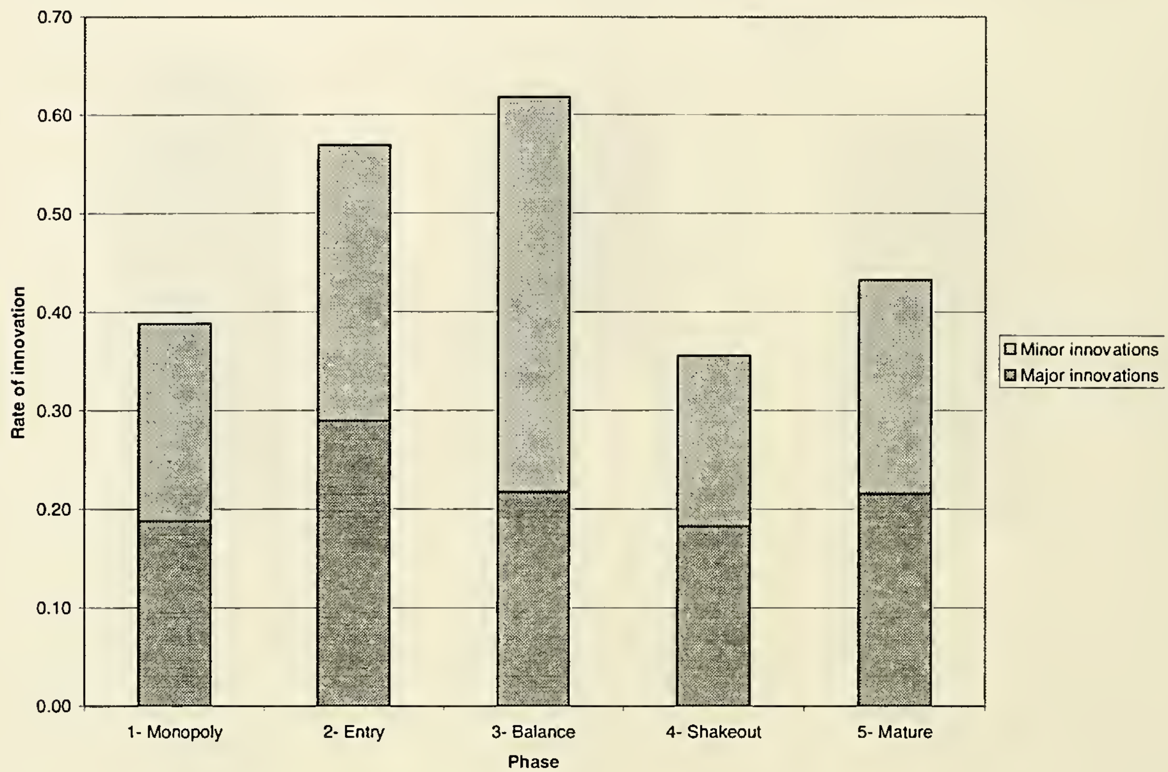
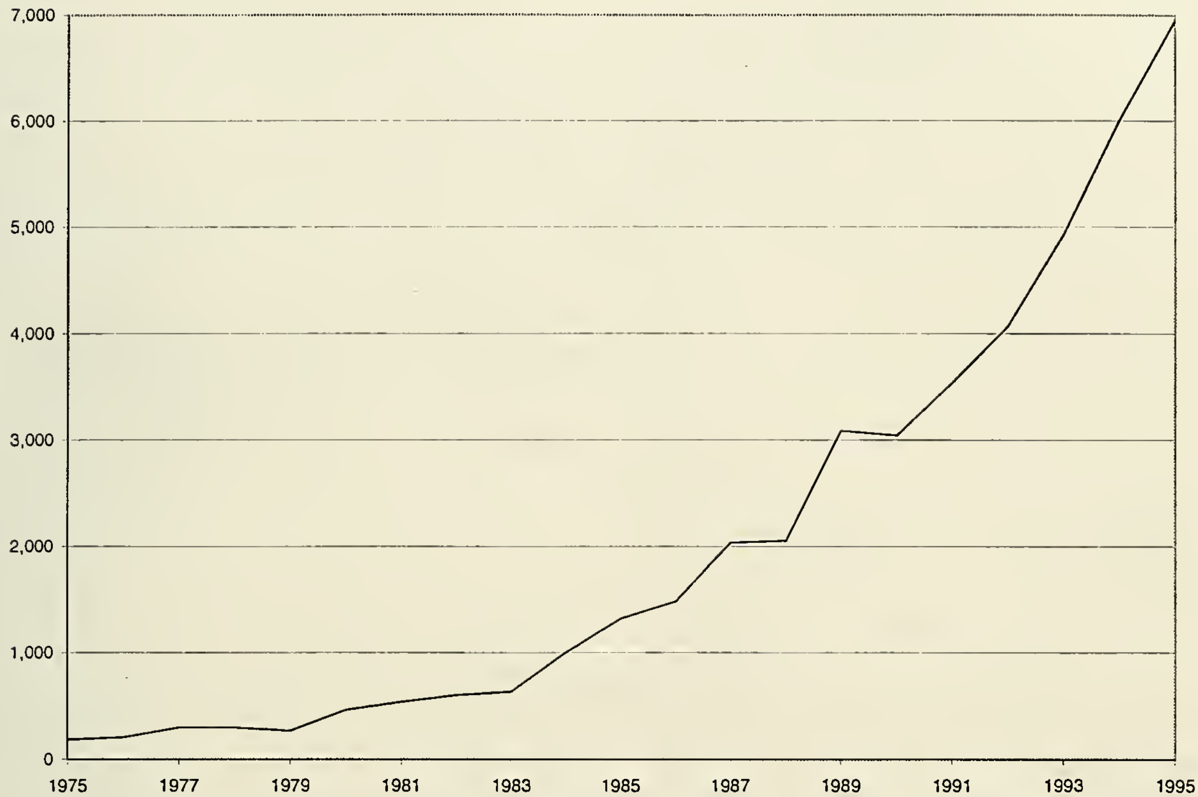
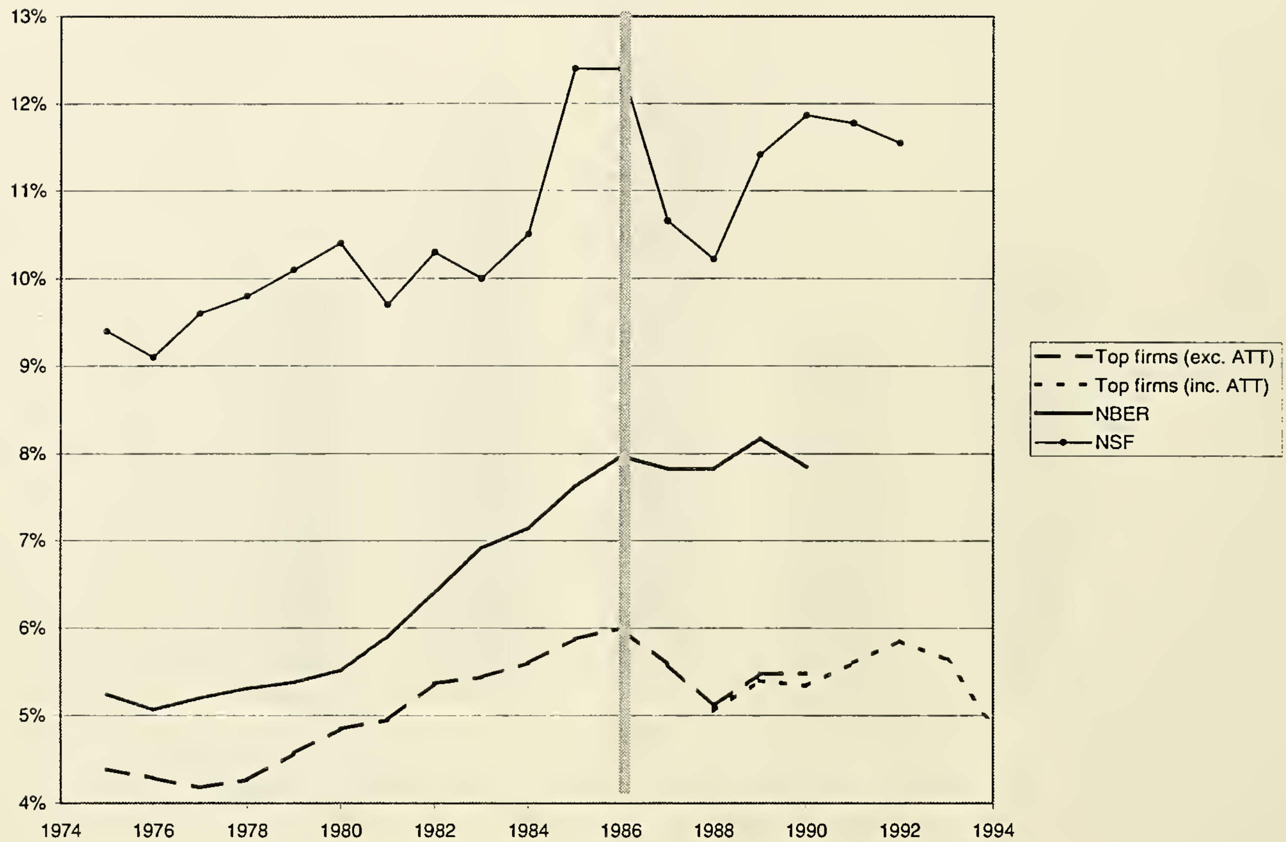


Figure 4 . U. S. Software Patents Issued



Source: Internet PATNEWS service.

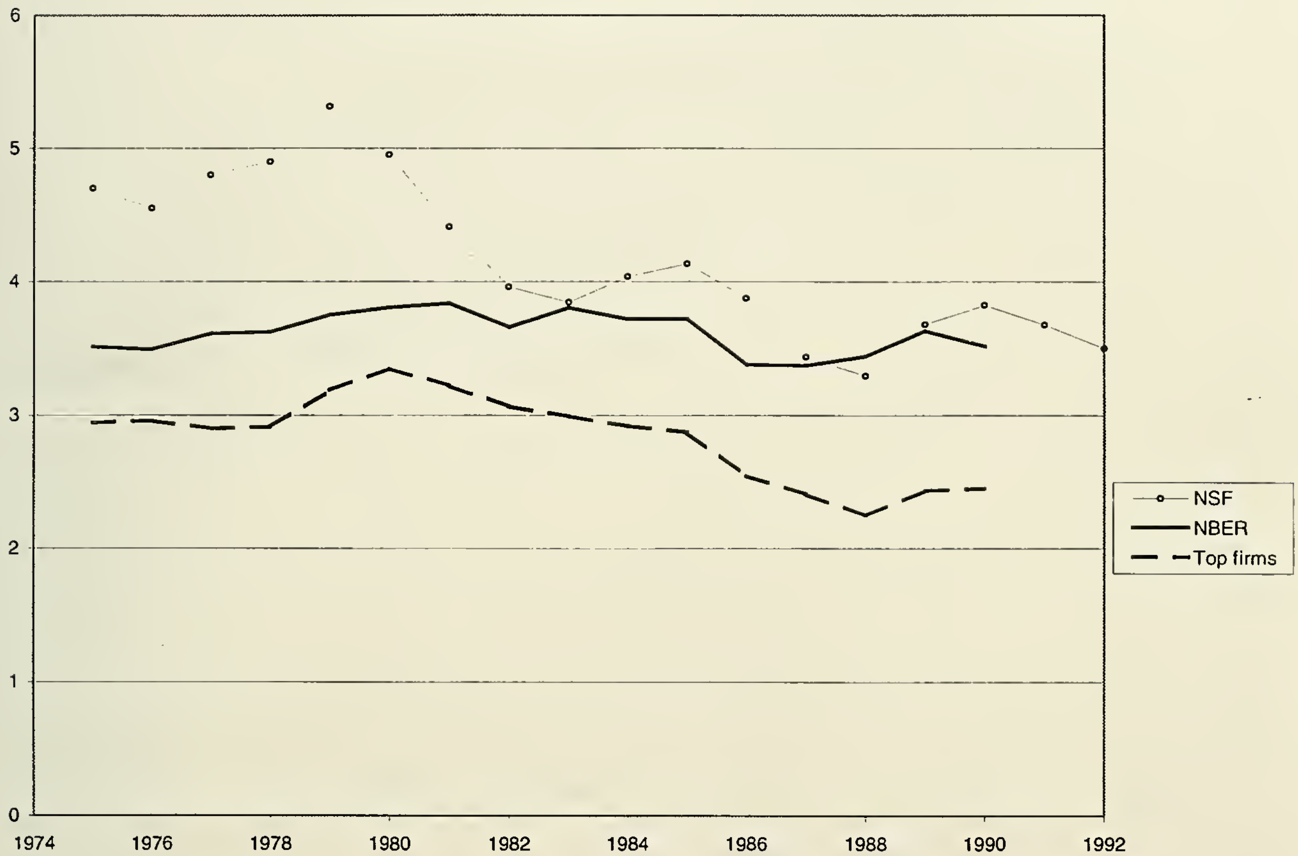
Figure 5. R&D Intensity for Software-related industries and firms



Sources: NSF *Research and Development in Industry, Science and Engineering Indicators*
NBER R&D Masterfile, Annual reports.

NBER series includes SIC 357, 365, 366, 367. NSF series includes SIC 357, and after 1986 part 737 and part 871. NSF series includes sample changes and hence is not directly comparable from year to year. Top firms come from Patent News Service rankings of software patents issued in 1995.

Figure 6. Relative R&D Intensity

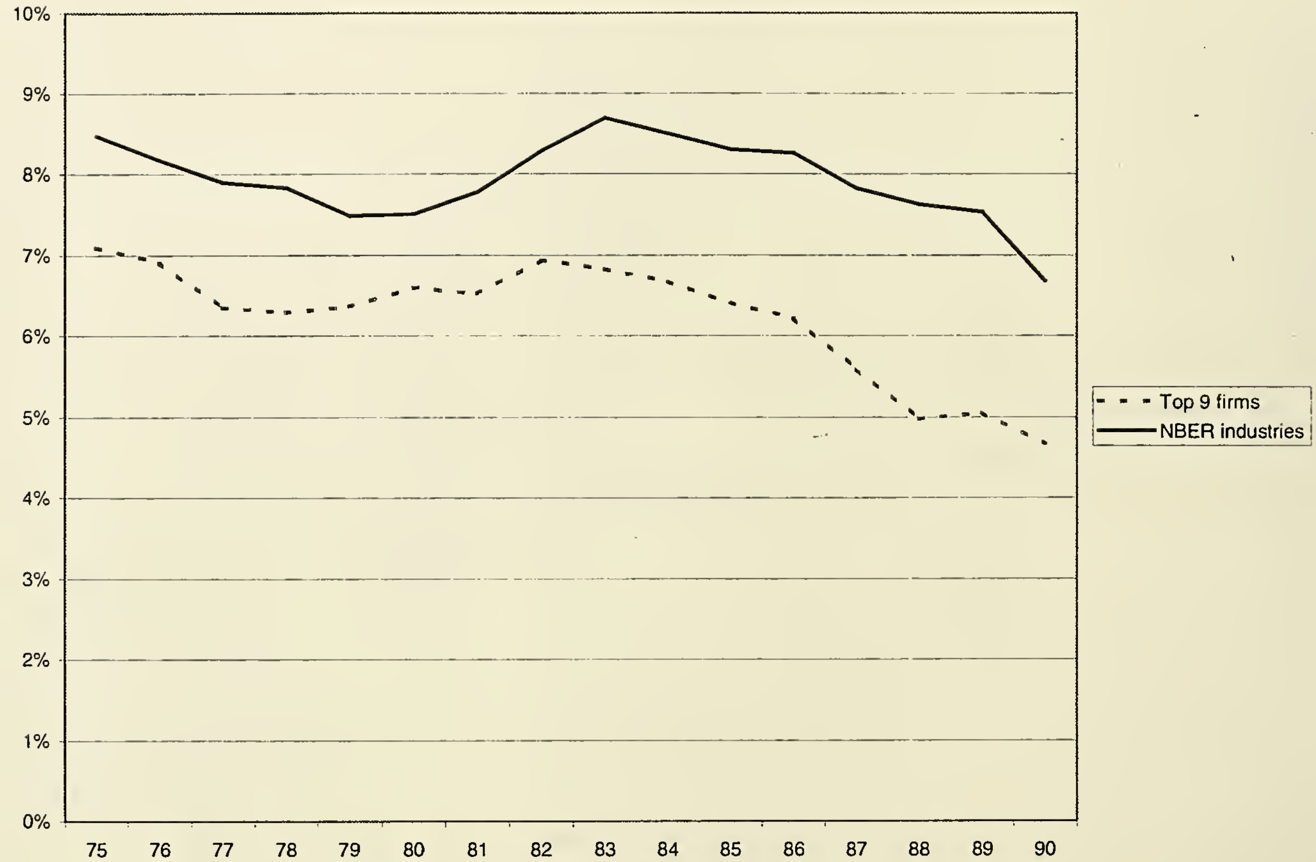


Sources: NSF *Research and Development in Industry, Science and Engineering Indicators*
NBER R&D Masterfile, Annual reports.

Relative intensity is the ratio of R&D spending to output divided by that ratio of R&D spending to output for the entire manufacturing sector. NBER series includes SIC 357, 365, 366, 367.

NSF series includes SIC 357, and after 1986 part 737 and part 871.
Top firms come from Patent News Service rankings of software patents issued in 1995 relative to the NBER series for all manufacturing.

Figure 7. Real R&D / Real Output (Deflated R&D Intensity)



Sources: NBER R&D Masterfile, Annual reports.

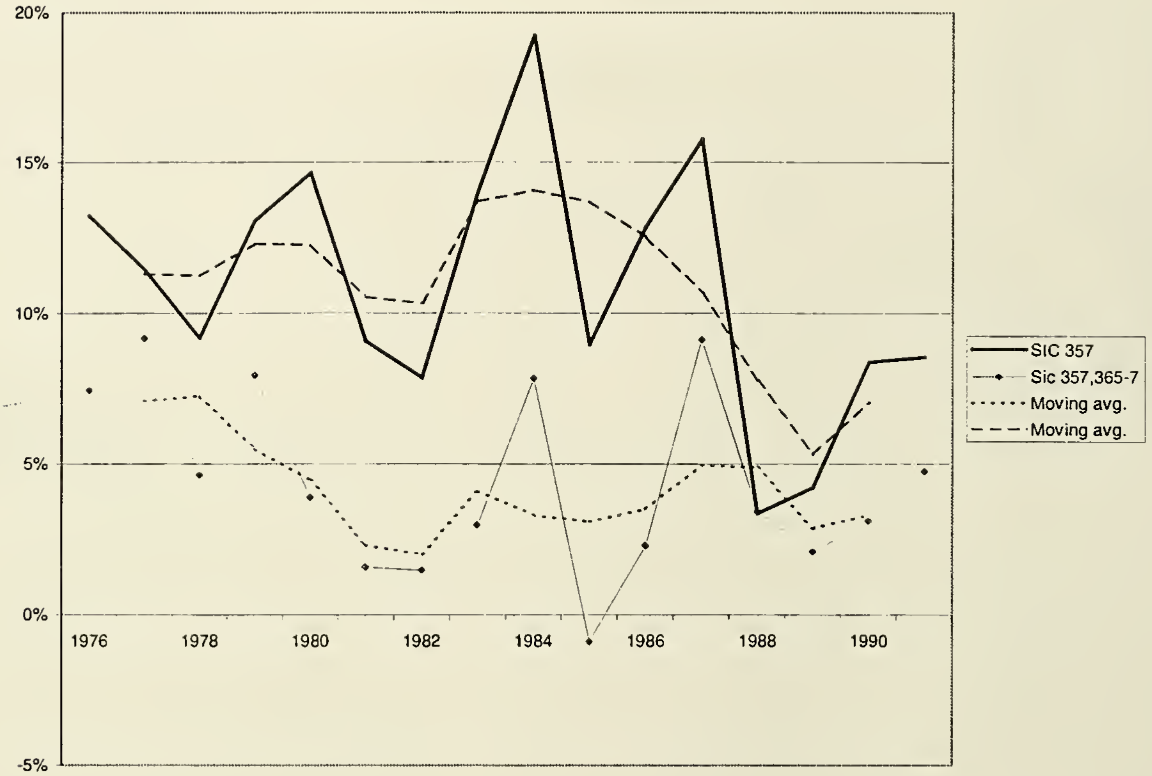
NBER series includes SIC 357, 365, 366, 367. R&D is deflated using NBER R&D deflator. Sales are deflated using a shipments-weighted mean for these industries from the NBER Productivity Database.

Figure 8. R&D Intensity for Small and New Firms

Source: NBER R&D Master File, Annual Reports

"Small firms" is a balanced panel of 49 firms from the software-related industries found in the NBER Master File in both 1980 and 1990 that had fewer than 1,000 employees in 1980. "New firms" are firms from the NBER Master File in software-related industries that first appear after 1973 and that had fewer than 5,000 employees their first recorded year. The unbalanced panel of new firms includes only the first eight years that a firm appears in the file.

Figure 9. Total Factor Productivity Growth of Software-related Manufacturing Industries



Source: NBER Manufacturing Productivity Database, Output-weighted aggregates of Bartelsman & Gray [1996] calculations for 4-digit industries. Moving averages are over three years.